TRAPS AND ERRORS IN RISK ANALYSIS

Lecture at George Washington University
Washington, DC
Richard Wilson
Department of Physics
and
Energy and Environmental Policy Center
Harvard University
Cambridge, MA 02138

March 31, 1982

Risk analysis in its most general form is commonsense combined with arithmetic. The errors are errors of commonsense. Rarely do we find errors of arithmetic, only failures to do the arithmetic. The fact that traps and errors exist should not, in my view, lead us to abandon risk analysis, but on the contrary, emphasizes the need for good risk analyses.

The first, and most often discussed error, is the demand for zero risk which appears in various forms. But while it is the new wisdom to regard the zero risk advocates as rabble rousers, or at best middle headed idealists, I want to emphasize that there has in the past been a role for an attempt to reduce the risk to zero. This was the case when the major risk of premature death was that of communicable disease.

This is brought out most clearly in a paper by Sir Richard Doll.[1] Doll showed that by most simple measures, health is better now than it was 100 years ago and is improving. This is because there has been a steady reduction in disease. I can also

[1]Doll, Sir Richard, "The pattern of disease in the post-infection era: national trends," Proc. R. Soc. Lond. B205, 47-61, 1979.

illustrate this by a plot of death rates versus time (Figure 1). The death rates are decreasing in all age groups except the one (15-24) where teenagers kill themselves with automobiles.

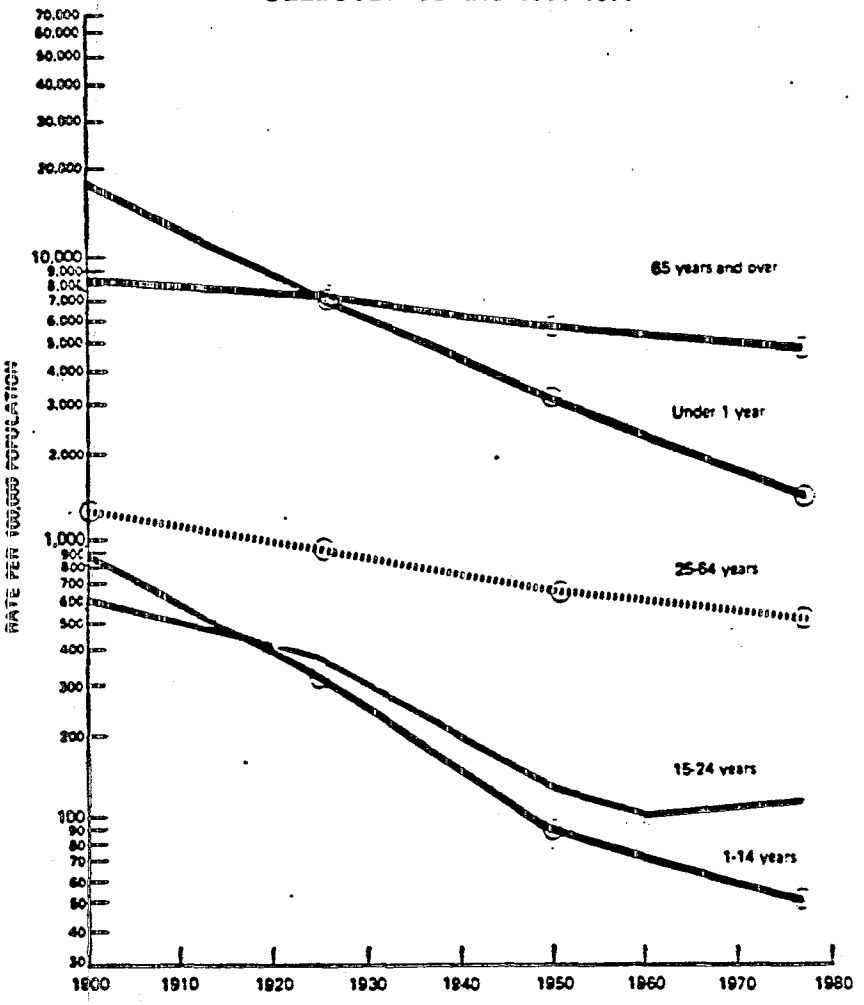
We can also go back over a longer period and look at life expectancy. In the year 800 A.D. life expectancy was about 28. It increased to 45 one hundred and fifty years ago, and is now 72 years for men and 76 for women (I personally am jealous of the women). The number of people surviving to a given age falls fairly slowly (Figure 2, but falls off sharply at age 70; an interesting age, because according to the nineteenth psalm, "the days of our years are three score years and ten" and nowadays compulsory retirement in the U.S. Does not begin until then. Death before age 70 can be considered premature and formerly premature death was mostly from diseas

Diseases were removed as causes of death, not by medical advances, but by technical and scientific ones. An Oxford historian has pointed out the influence of the drainpipe (for main drainage) in the European history of the early 19th century. Chlorination of drinking water also played a major role. We must not, therefore, turn our backs on technology, but use it wisely. But there is an important corollary; the expense of reducing disease was moderate.

In each case there were a large number of dead bodies obviously attributable to a disease; and once the cause was recognized, a massive and rapid technological effort was made to remove the hazar completely. This led to a general view that once a risk of disease is recognized it is possible to complete eleminate it--and it shoul be completely eliminated. This, then, is a justification for the

FIGURE 1

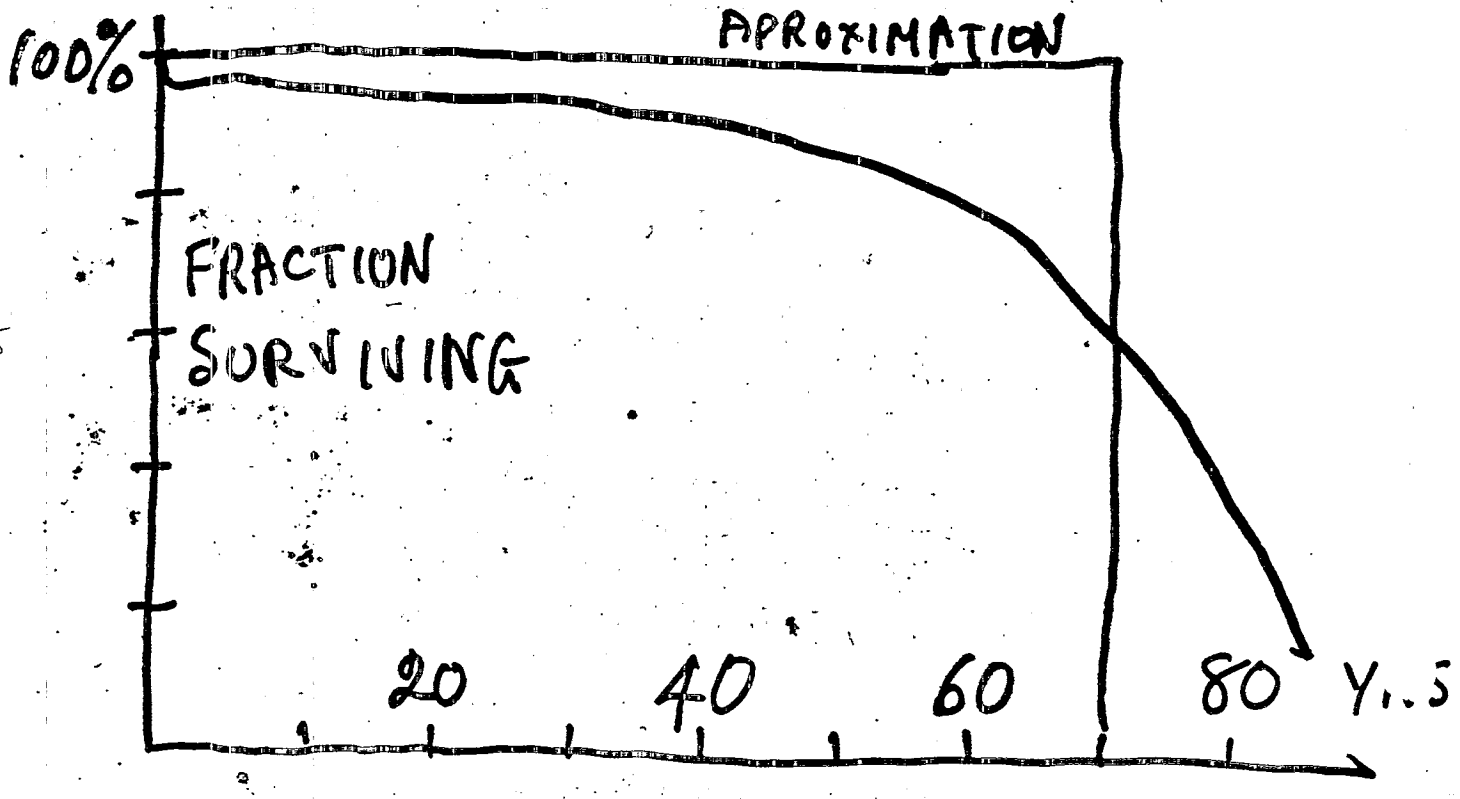
DEATH RATES BY AGE: UNITED STATES, SELECTED YEARS 1900-1977



NOTE: 1977 data are provisional; data for all other years are final. Selected years are 1900, 1925, 1950, 1960 (for age group 15-24 years only), and 1977.

SOURCE: National Center for Health Statistics, Division of Vital Statistics. (HEW)

2025546325



zero risk approach.

The reason that the zero risk advocates are wrong is that the situation has changed dramatically. We have almost removed communicable disease in the Western world as a cause of death.

The second important error in a discussion of risks is a failure to recognize the changed nature of the risks which terminate our lives. The earlier risks can be called historical; the risk could often be calculated from historical data, and once the idea was suggested, the proof of causality was straightforward. There are some people, and for politeness I will not refer directly to them, who have called for only regulating on these risks; for regulating only known measureable effects on health. If taken literally, this would prevent our controlling a large number of risks which are well worth reducing. We must recognize that risk is an expression of uncertainty. The uncertainty can arise in two ways; we can for example believe that cars kill people who cross the road, yet be uncertain that an individual is killed; or we can be unsure whether a new technology will ever kill anyone or not. The new risks of society are for new technology, about which we must speculate. It is improper to say "we do not know if there is a risk of air pollution at present levels." If we do not know--there is a risk, but the magnitude of the risk may be uncertain.

These two errors are at the root of our present discussion about the Clean Air Acts. On the one hand some people still want zero risk, forgetting that air pollution is not a communicable

disease and zero risk makes no sense, and on the other hand others are unwilling to regulate until dead bodies can be counted and reliably attributed to the source.

No risk analysis is perfect. In those I have reviewed, and Crouch and I discuss a number in our book on Risk/Benefit Analysis[2], the most important is omission of an important risk. It is foolish to compare risks of nuclear power and coal without mentioning the risk that either may contribute to war; the one perhaps by making nuclear fuel more easily available, the other by making energy more costly. We must insist on putting numbers on each and every risk we can, but there are some risks that cannot be easily expressed in numerical terms. If the numerically expressed risks are properly discussed, these other risks will be highlighted; all too often they are ignored. For example, a 400 page report on LNG safety ignores questions of sabotage. Instead of saying so in the abstract, introduction and summary, the sentence is hidden in the middle. Yet for LNG and LPG facilities it is my considered view that the siting should be governed by the possibility of sabotage.[3]

Samuel Johnson once wrote that "Round numbers are always false" and a risk analysis without a discussion of uncertainty can be false. As I noted before, it is in the region of uncertain risks--risks of a consequence which we are not even sure exists--

---

[2]E.A.C. Crouch and Richard Wilson, Risk/Benefit Analysis, Ballinger Press, Cambridge, MA, 1982.

[3]Richard Wilson, Testimony to Energy Facilities Siting Council, Massachusetts (1979).

that a risk analysis is the most important. In this case, however, it is vital to state the assumptions--early and often, as it used to be said about instructions to Cook County voters. But it must not be thought that an uncertain number is useless. Dr. Samuel Epstein and Samuel C. Florman in two separate articles in the Technology Review[4, 5] argue that the estimates of risk of exposure to vinyl chloride at low doses vary by a factor of 100,000, and are therefore useless. They are wrong on both counts. Estimates vary by an infinite factor; some scientists argue that the risk is finite, others that it is zero. Any finite number divided by zero gives infinity (not 100,000).

Yet risk analysis can be useful; as of 1981, 82 cases of liver angiosarcoma had been attributed to occupational exposure of vinyl chloride world wide. These were due to past high exposures over a 20 year period. Liver angiosarcoma is a rare cancer, and few attributable cases would have been missed. Allowing for some cancers still in the latency period, cancers at other sites in the body, and cancers from exposures in the general environment, I find a maximum of 1000 over 20 years from past exposures. Exposures have now been reduced 1000 fold, and most scientists would agree that the number of cases will fall at least linearly to 1 every 20 years worldwide, and I believe no liver angiosarcoma has been attributed to exposures in the last 5 years.

[4]Samuel Epstein, "Cancer inflation and the failure to regulate," Technology Review, Dec./Jan., 1980.

[5]Samuel C. Florham, "Living with technology: tradeoffs in paradise," Technology Review, Aug./Sept., 1981.

A number of people argue that natural foods are better and our problems of cancer and so on arise from artificial chemicals. It is conventional wisdom to ridicule this. This ridicule is illustrated in a New Yorker cartoon (Figure 3) showing that natural ingredients were used for witches' brews as well as for honest men's sustinence. Indeed, I like to point out to our local health food store that peanuts dried in the sun dry more slowly than those dried artificially and develop more molds and therefore have higher concentrations of aflatoxin B1. But it would be an error to ridicule the "health food nuts" without trying to see if there is a legitimate point in their favor. There is.

We have historical data on natural foods, and believe a major disaster is unlikely, whereas if a new chemical (such as thalidomide) is put on the market, thousands of people may be hurt before we know what the danger is because of a latent period.

But this is no excuse for turning our back on technology. Natural phenomena may wipe out the human race, and got close to doing so in the black death. That particular risk is probably now eliminated by technology.

It is important to understand a proper flow of information in reaching decisions about risk. Risk analysis is an aid to a decision and should not be a decision in itself. It is a grevious error to confuse the risk analysis with the decision itself.

"It's going to be great! All natural ingredients."

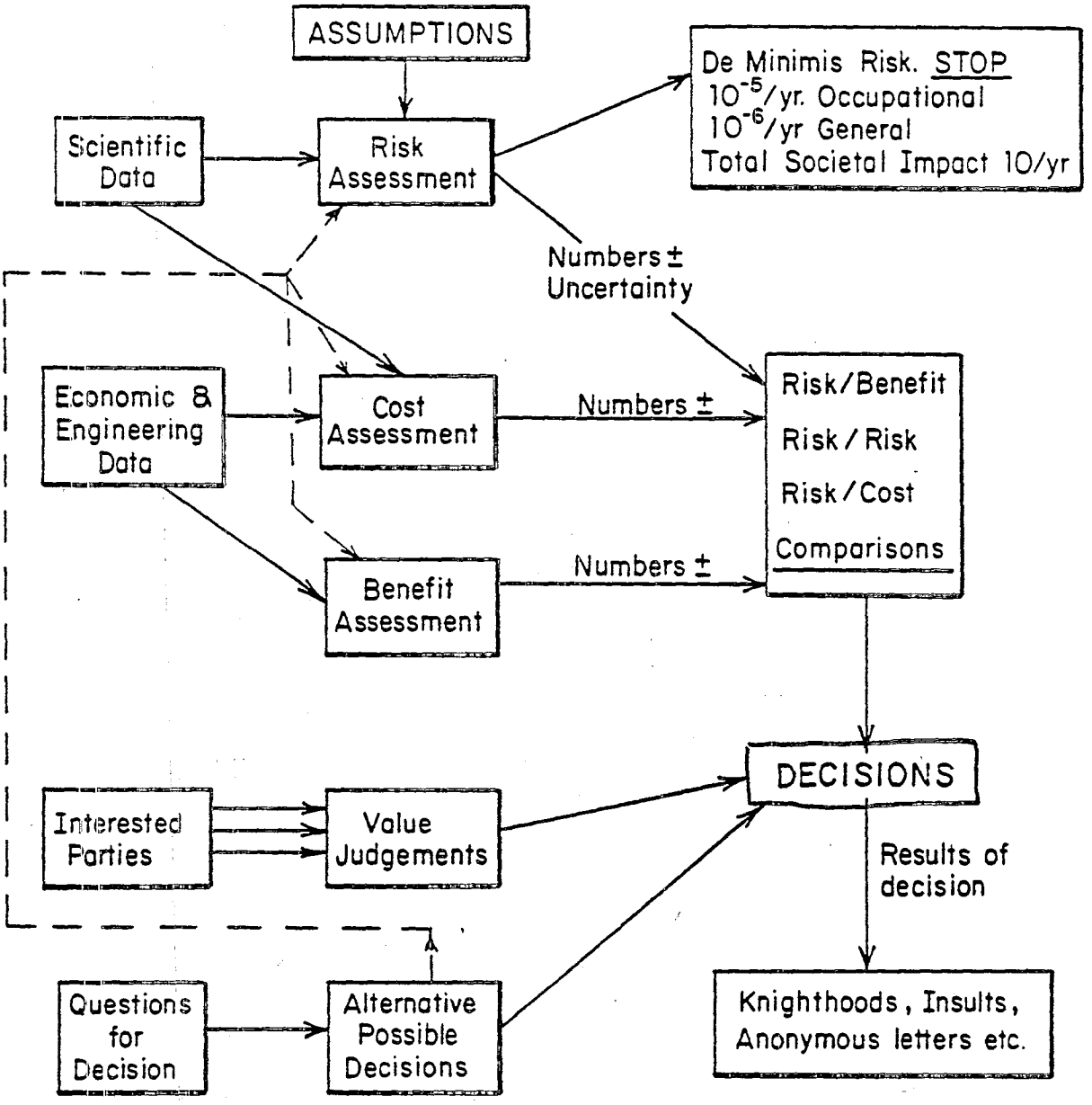
I illustrate the flow of a discussion of risks in Figure 4. The information flows from the scientist to the assessor; to a comparison with costs and benefits to a decision maker (evaluator) who may be a bureaucrat, a politician or an ordinary citizen. All interested parties, unions, public, churches, industry, fishermen, etc. must be identified and these will have their own value judgments. Value judgments enter into the decision, but should not influence the assessment and comparison. Of course the assessor has to know the questions for decision, and some idea of the alternatives that are open to the decision maker in order to write down his assessment in such a way that the decision is properly illuminated. For this reason I put a dotted line between questions for decision and assessment.

The risk assessment is most useful when it is kept distinct in this way. The Carcinogen Assessment Group of EPA has been steadily improving in this regard; they usually quote the upper 95th percentile of the risk, but state also that it might be zero. They nowadays avoid directly supporting a decision.

I make here an analogy with the Anglo-Saxon legal system. The judge decides on questions of law; the jury--the assessors--decide questions of fact. The judge instructs the jury before they assess the facts as to what is legally relevant.

However the separation is sometimes deliberately broken as a historical example shows. In Cromwell's time, a law was passed making it illegal to kiss in public. Perhaps the law was sensible,

RISK ANALYSIS



but there was a draconian penalty---death. However all juries violated the rules, and no jury would convict anyone accused of the crime.

Similarly if a group of assessors are asked an inappropriate question, or suspect that their answer may lead to stupid draconian measures, they may well decide not to state the true opinion. Thus if the question is merely "is this chemical a carcinogen" with no question of potency, I believe it is now inappropriate. If it is coupled with a draconian action of a complete unquestioning ban if the answer is yes, many scientists will refuse to play the game.

We can enter this diagram at any point; but usually it will be the decision maker who asks for an assessment. We must also be aware that if well done, a risk assessment may illuminate several different decisions. For example, the risk assessment for benzene may apply either in the workplace, or for the environment.

Errors often arise from not realizing that there are different constituencies in risk decisions and each constituency must be satisfied. It is an error to calculate the average risk to society, show that it is small and ignore it in spite of the large risk to an individual segment. I illustrate this by calculating the risk of being killed by a bear in Glacial National Park. According to the figures available to me, and as a risk assessor I accept this data, 2 people were killed in 1967; 1 person in 1975; and 3 people in 1980. That amounts to 6 people over about 15 years, or 0.4 per year. The Park Service informs

me that the total number of park visitors is 1,500,000/yr; about 18,000 backpack campers/yr (25,000 permit nights); and about 1000 employees of park concessions. Of the 6 people killed, all were backpackers, and 4 were concession employees.*

The risk averaged over the whole U.S. population is:

$$\frac{6}{15} \times \frac{1}{220,000,000} = 2 \times 10^{-9} \text{ per year.}$$

I do not expect the President of the U.S. to lose sleep over it. It is a de minimis risk.

The risk averaged over park visitors is:

$$\frac{6}{15} \times \frac{1}{1,500,000} = 3 \times 10^{-7} \text{ per year;}$$

still de minimis by my criterion, but on the border of worth bothering about and indeed the park service have done little.

The risk as a fraction of backpack campers is:

$$\frac{6}{15} \times \frac{1}{18,000} = 2 \times 10^{-5} \text{ per year.}$$

This is a voluntary risk worth reducing, but probably acceptable and, indeed, my wife and I accepted it a few summers ago (but we did take care; put bells on our packs and hung the food on poles).

The risk as a fraction of park concession employees is

$$\frac{4}{15} \times \frac{1}{1,000} = 3 \times 10^{-4} \text{ per year.}$$

This is now large, as large as the risk of driving a car to get to Glacier Park in the first place, and probably worth addressing by the representatives of this smaller group.

But this analysis leads, if we are not careful, to another trap. Why do you include in the denominator those who are not

*This is based on oral statements only.

exposed? Because we may know no better. But we should go on reducing the size of the denominator until one reaches a logical trap. We should include in the denominator all those "at risk" and not only those who were killed. We get 6/6 or a risk of unity when we identify the group at risk as the group actually sleeping in the path of the bears and subsequently mauled. The denominator must be chosen so that the risk calculation has predictive power. To restrict the denominator too much takes away that power.

Bross, of the Rosswell Memorial Hospital in Buffalo, fell into this trap when discussing leukemia incidence. He restricted the group at risk to those who already had health problems which are precursers of leukemia. As shown by Bond[6] his calculations have no predictive power.

Another error is to assume that a risk benefit analysis should be constant for all time. This omits the possibility of technology improvement. It is clear that the public wants risk to be reduced, and will not stand for additional risks being introduced to society--unless there is an obvious compensating risk reduction.

Therefore we must constantly search for ways of improving technology. I illustrate this by two examples.

From the earliest days it was known that x-rays cause skin cancer and in the 1920's it was found that radiation causes leukemia and in the 1930's and 1940's it was found that radiation causes a host of other cancers too.

[6] V. Bond, BNL report.

In the 1920's physicians pointed out the huge benefits of rapid diagnosis by use of x-rays, and said that the benefits outweight any conceivable risk due to the x-rays. Of course they were right by a reasonable standard. This risk benefit comparison was particularly simple because the items were measured in the same units. X-rays save lives and these can be compared directly with lives at risk from the radiation induced cancer.

But the overall comparison of risk and benefit and the assurance that benefit exceeds risk is only the first question in a risk benefit approach. In the 1920's it was pointed out that the same benefit of rapid diagnosis by use of x-rays can be achieved with less risk: more sensitive x-ray film; image intensifiers; shielding against stray x-rays; and indeed it was not very expensive to do so. But these risks have only slowly been reduced, from many rads per x-ray in 1920 to 1 rad in 1950 to 5 millirads today. Why was the reduction slow?

Rachel Carson, in her much misquoted book, Silent Spring,[7] agreed that pesticides were important and that the benefits of their use outweigh the risks. Again, if we think carefully, we can put at least some of the benefits in the same units as the risks; fewer pests means more food; more food means better nutrition; better nutrition reduces health risks. But Rachel Carson insisted it should be possible to have the same benefits with less risk by more careful use.

How do we ensure that we constantly address this question?

[7] Carson, R., Silent Spring, Houghton Mifflin, Boston, 1962.

I believe that we must have a continuing decision process. The procedure in each of the above examples was to show that benefit > risk and the decision therefore was that the action should proceed. But in all cases of new technologies it seems appropriate to look further and ask whether we can reduce the risk with a modest cost.

In setting the air quality standards for moving sources (automobiles), Congress demanded a technology forcing approach. The standards were set progressively lower in the future and it was up to the automobile industry to find the technology to meet them. The industry screamed, but a Committee of the National Academy of Sciences chaired by Dr. Edward Ginzton of Varian Associates agreed that it was technically possible to meet this standard, and they are being met.

Is this technology forcing procedure the best way for technology improvement? Congress, after this success, thinks so. If we feel otherwise we must find a better way. I throw on the table for consideration the following idea. When any new product or process is accepted, there must be a fraction (1 to 10%) of the profits spent on a study of the alternatives to the actual action; other pesticides; less use (including how to restrain overzealous salesmen); better x-ray machines--and less x-rays, including how to prevent x-rays being taken merely to protect against liability in a negligence suit. This money could be paid by industry to NAS or a foundation to do this study.

I now bring to your attention four separate court and regulatory actions that recognize the importance of risk assessment and its more general extension, risk analysis.

1. The Court of Appeals,[8] in the case where FDA questioned the safety of Monsanto's plastic bottles made with an acrylonitrile polymer, stated that the administrator of FDA can ignore de minimis risks, notwithstanding the superficial rigidity of the Delaney Clause.

2. The U.S. Supreme Court decided in the benzene case[9] that the Secretary of Labor for OSHA had to find (technical/legal term) that there is a significant health risk before proceeding and by quoting the one and only risk assessment before the court implied approval of risk assessment as a procedure for deciding whether a risk is insignificant or not. Chief Justice Burger stated that "inherent in this statutory scheme is the authority to refrain from regulation of insignificant or de minimis risks."

Although the Court quoted my testimony in this case, they failed to quote my successful effort at risk reduction. At the public hearing, the administrative law judge, and half the audience, were smoking and the room was filled with haze. I objected to being exposed to such dangerous carcinogens in the workplace. Now the hearing room is covered with numerous, large "No Smoking" signs.

---

[8]Monsanto, et al., v. Kennedy, FDAO, 613 F 2nd 947 (DC Circuit)), 1979.

[9]American Petroleum Institute v. OSHA, et al., 581 F 2nd (DC Circuit), 1978; 100 S. Ct. 2844 (1980).

3. Now we have a consent order[10] between FDA and several cosmetic manufacturers. In this "FDA agrees to propose the utilization of scientifically accepted procedures of risk assessment and to raise the issue as to whether, in the view of the procedures, hair dyes containing 4-MMPD present a generally recognized risk to human health."

4. The FDA made a similar decision in deciding that the risk of lead acetate hair dyes (the Grecian formula) is insignificant.

As I look at a 5th case, the cotton dust case, it appears to be different. It has been argued that the Supreme Court decision is a repudiation of risk analysis. But we must note that "exposure to cotton dust represents a significant health hazard to employees."[11] The health hazard at current levels, acute byssinosis although reversible, is directly observable, unlike the effects of benzene at low doses which is merely inferred by extrapolation. I calculate with a very rough first guess using the numbers in reference 11 that the cotton dust standard proposed by OSHA saves cases of byssinosis (brown lung disease) and hence premature death at a cost of between $50,000 and $500,000 per case--less than the $1 million per case often spent to reduce occupational disease, and $10,000,000 to reduce occupational death and much less than the $240,000,000 per calculated case I found in my calculations in the benzene case. Of course, we would like to understand the connection (if any) between

---

[10] Carson Products, et al., v. DHHS, et al., Civil Action No. CV 480-71, U.S. District Court for the Southern District of Georgia, Sept. 1980.

[11] 101 Supreme Court Reporter, p. 2478, American Textile Manufacturers v. Donovan and National Cotton Council of America v. Donovan, 43 Fed. Reg., 27350, col. 1.

acute byssinosis and life shortening effects. Data on this is scarce and represents a challenge for risk analysts.

My cheerful tentative conclusion is that the Supreme Court behaved with commonsense in both decisions. If OSHA had perfomed a reasonable risk-benefit analysis it is not obvious that they would not have reached the same decision, or decided to lower the standard for cotton dust even lower than 200 $\mu g/m^3$ until they matched $1,000,000 to $10,000,000 per life--unless of course they found other, more serious secondary effects on the economy.

The point here is that, as far as I can tell, there was no widely accepted risk analysis presented in this case; and OSHA's decision to take prompt action to tighten the standard is then not unreasonable.

I now come to several errors that arise from inadequate comparisons of risks. For cancer risk calculations, Crouch and I[12] have been emphasizing the importance of including uncertainties in a risk calculation. I illustrate this by going over an argument I had with EPA's director of water quality standards at the Toxicology Forum in February 1981. I objected to his setting of a standard of 100 ppb for all halomethanes; drinking 2 1/2 liters of water containing chloroform every day for life can be shown to give a risk of 4×10^{-5} per year or 3×10^{-4} per life. This is based on several assumtions; that the dose response relationship is linear; that men in their lifetimes have as great a risk of cancer as animals in their lifetimes, when exposed daily to the pollutant

[12] Edmund A.C. Crouch and Richard WIlson, "Interspecies Comparison of Carcinogenic Potency," Journal of Toxic. and Environ. Health, 1979. Edmund A.C. Crouch and Richard Wilson, Journal of Risk Analysis, 1, 1 (1981).

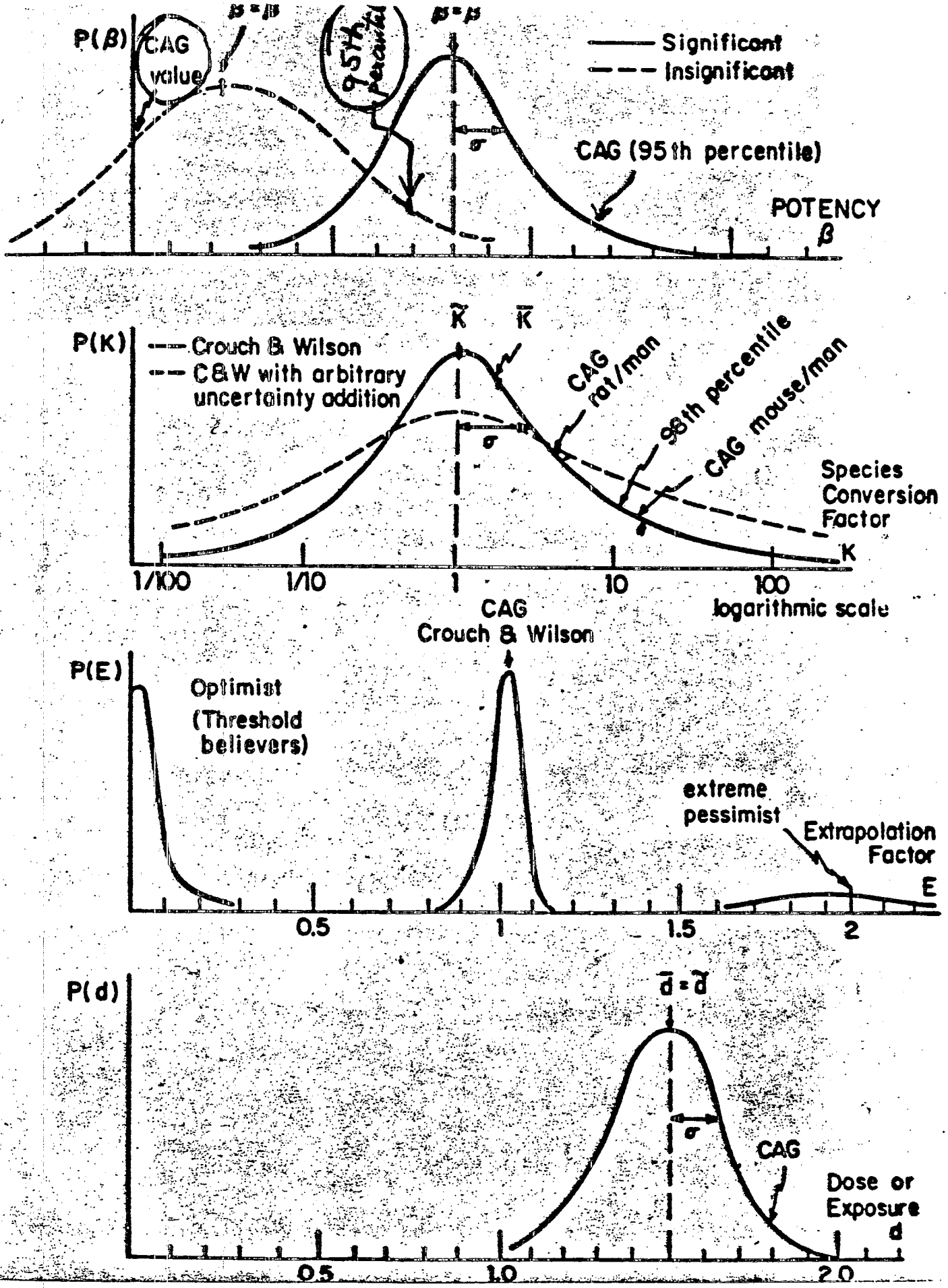
with the same fraction of body weight.

EPA have the same standard for all halomethanes. Yet bromomethanes are more toxic and more mutagenic; by skin painting tests more carcinogenic. The exact amount is uncertain, but it is hard to see how the risk could fail to be greater at equivalent exposures, and therefore the standard should be lower.[13] Similarly, tribhlorethelene and perchlorethelene are less toxic, and less carcinogenic (if carcinogenic at all) and the allowable exposure might reasonably be greater--or at least it should be explained why it is not.

We can reintroduce the zero risk error in a new form if we demand too low a risk, and demand too conservative a way of calculating it. FDA has proposed[14] in a discussion of accidental food additives such as DES, that risks less than 10^{-6} per lifetime (1.4×10^{-8} per year) be ignored as de minimis and by implication other risks should be regulated. This is a very low number. Moreover there are very conservative rules proposed for calculation. By simple extension of these rules, I would find that many municipal drinking water systems in the U.S. have risks of 10^{-4} per year. It is clear that the FDA rule cannot be applied to all situations, and it therefore needs more discussion to show just where it should be applied (if at all). At the present time, most agencies are inconsistent without saying so.

In the next two figures I show how this comes about. The risk, according to our approach, has 4 factors:

14



$R = \beta d K E$

β is the carcinogenic potency defined as the slope of the relation connecting tumor incidence and dose;

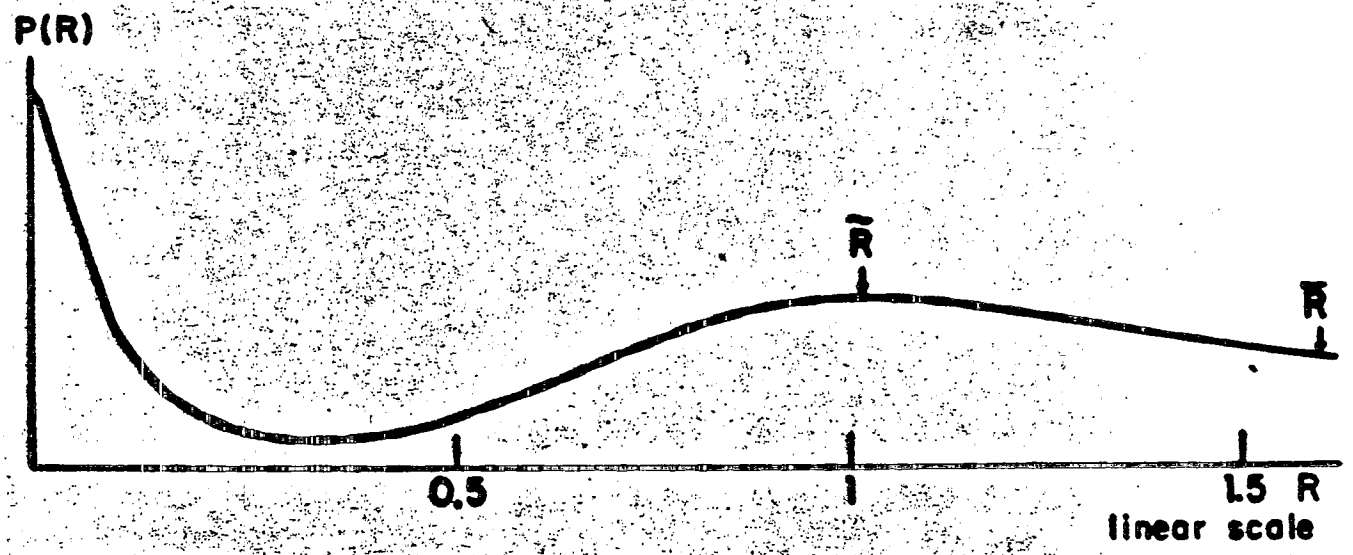
d is the dose (or exposure);

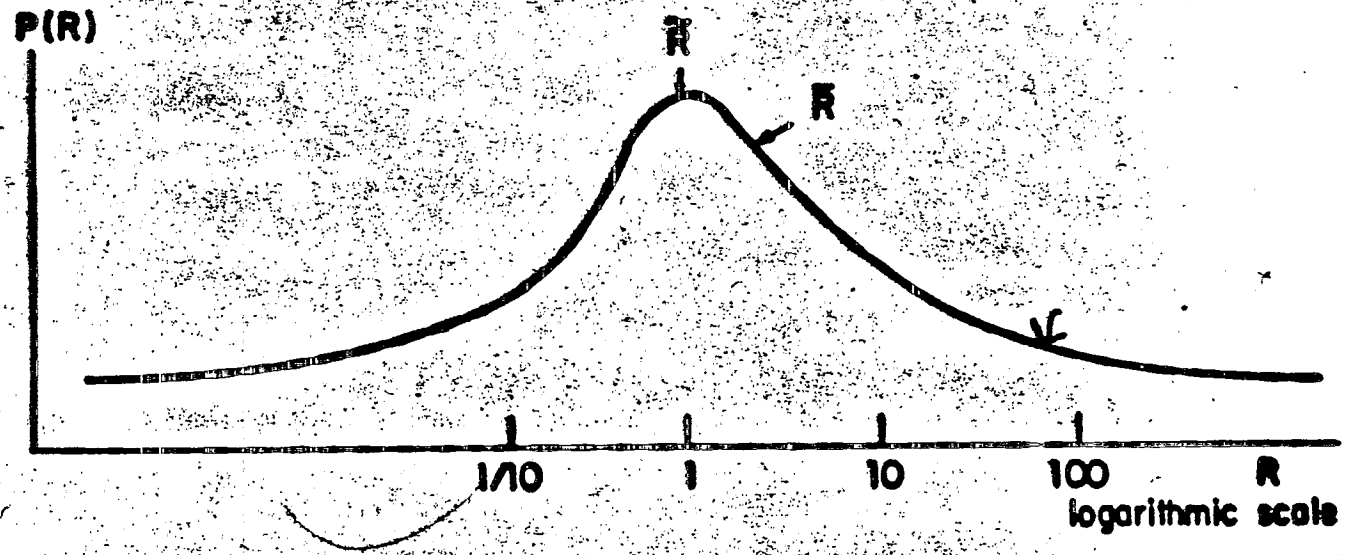
K is an interspecies conversion factor. We take a probability distribution for this, centered around unity when animal and human risks are compared for equal lifetimes and doses are compared as a fraction of body weight.

E is an extrapolation factor; to relate the risk at high doses (measured in animals) to that at low doses. E = 1 in the present conventional wisdom, but if we believe in a threshold we can put E = 0.

In Figure 5 I show a calculation of β (in animals) from data, and the spread of the probability distribution $P(\beta)$ around the best value $\tilde{\beta}$ is given by statistics. The solid line is when the chemical is identified as a carcinogen, the dotted line when it is not because $\tilde{\beta}$ is not 3 standard deviations from zero. CAG take the 95th percentile point as shown for the solid line, but take 0 for the dotted line. We believe this is inconsistent. It leads to regulatory hiatus, improper signals to industry, and improper comparison of risks.

We see this by noticing that the data for saccharin looked like the dotted line before 1977 and that for formaldehyde before 1981. Neither chemical was regulated. Suddenly, on April 15, 1977, the FDA declared saccharin that is a carcinogen when new data became





available, and when in 1981 unpublished data from an NYU study became available, formaldehyde became a carcinogen and strongly regulated.

In each case there were prior indications of probable carcinogenicity. The regulatory hiatus would have been avoided if both chemicals had been regulated on the basis of the 95th percentile of the risk even before the experiments became sensitive enough to have statistical significance. Moreover, under the present CAG procedures, it is to industry's benefit not to perform a sensitive experiment, and change a chemical's status from no regulation to stringent regulation. Under our suggestion, the chemicals would be regulated in both cases, and shrinking the statistical error might actually reduce the impact of the regulations.

For K, we take a broad probability distribution P(K), as discussed in reference 12. CAG take a fixed number for K, with no uncertainty, but since they compare animal to man on a surface area, rather than a weight basis, their number is close to the 95th percentile of our distribution as shown in the Figure.

The probability distribution for d presents no problems.

The combined distributions for the risk P(R) is presented in Figure 6; the top curve is drawn to a linear scale, the bottom curve to a logarithmic scale.

Now I distinguish between an accurate assessment, to calculate $\tilde{R}$ or perhaps the average risk $\bar{R}$, and a prudent assessment. A federal agency, such as EPA, has an obligation to protect the public, and will choose the upper 95th or 98th percentile of the distribution. This is proper. But in comparing risks, such as replacing

one pesticide by another, it would be improper to compare different moments of the risk distribution. If the 95th percentile is appropriate it should be calculated in both cases, including the case when the best value $\tilde{R}$ or the mean value might be zero, or less. I hope we can get this point accepted some time. If and when it is, a corollary I predict is that it will be found impossible to consistently regulate risks below 10^{-6} per year (7×10^{-5}/life).

A much harder point is to realize that there is a risk, which should be calculated, even when a medical effect is not found (I noted this already in my discussion of error 2). Logically this is similar to the case when a chemical is not found to be carcinogenic. I note that a chemical cannot be found to be non-carcinogenic; only that its carcinogenic potency be less than a specified amount. Many papers by distinguised authors are logically imprecise here, and are therefore confusing. Formally we can express this by saying that $\bar{R}$ may be 0, and $\tilde{R}$ may be 0 or even negative, but $R_{95\text{th percentile}}$ is in such cases finite and often large.

I conclude by pointing out that as when we look at examples of the errors and traps that we can fall into when doing a risk analysis, we can avoid falling into them; moreover, in each example it becomes clear what a reasonable decision might be; whether to act or not to act; and if to act, who must reduce and control the risk.

In our bear example, it was not the president of the U.S.; the Park Service can provide information; but it is individuals who must act and probably the park concession operators must educate

their staff.

The proof of the pudding is in the eating. If risk analysis can help society in making decisions about risk--and in particular in reducing risk at reasonable cost--then risks analysis is worthwhile. Otherwise it will be a useless intellectual exercise.